

RESEARCH

Open Access



Evaluation of different time domain peak models using extreme learning machine-based peak detection for EEG signal

Asrul Adam¹, Zuwairie Ibrahim², Norrima Mokhtar^{1*}, Mohd Ibrahim Shapiai³, Paul Cumming⁴ and Marizan Mubin¹

*Correspondence:
norrimamokhtar@um.edu.my

¹ Applied Control and Robotics (ACR) Laboratory, Department of Electrical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia
Full list of author information is available at the end of the article

Abstract

Various peak models have been introduced to detect and analyze peaks in the time domain analysis of electroencephalogram (EEG) signals. In general, peak model in the time domain analysis consists of a set of signal parameters, such as amplitude, width, and slope. Models including those proposed by Dumpala, Acir, Liu, and Dingle are routinely used to detect peaks in EEG signals acquired in clinical studies of epilepsy or eye blink. The optimal peak model is the most reliable peak detection performance in a particular application. A fair measure of performance of different models requires a common and unbiased platform. In this study, we evaluate the performance of the four different peak models using the extreme learning machine (ELM)-based peak detection algorithm. We found that the Dingle model gave the best performance, with 72 % accuracy in the analysis of real EEG data. Statistical analysis conferred that the Dingle model afforded significantly better mean testing accuracy than did the Acir and Liu models, which were in the range 37–52 %. Meanwhile, the Dingle model has no significant difference compared to Dumpala model.

Keywords: Extreme learning machines (ELM), Electroencephalogram (EEG), Peak detection algorithm, Peak model, Pattern recognition

Background

Peak detection algorithms are prominently used for event classification in various physiological signals such as in electroencephalograms (EEG) for diagnosing epilepsy (Acir 2005), photoplethysmograms (PPG) for monitoring heart rate (Elgendi et al. 2013), and in EEG (Adam et al. 2014b) or electrooculograms (EOG) in the particular application of tracking eye gaze events (Barea et al. 2012). In all of these common applications, peak detection is commonly the first step in signal processing. For example, semi-automatic diagnosis of epilepsy can be based on the frequency of peaks detected in the EEG recording during a given time interval. A similar approach is used for identifying eye blink events, a frequent source of interference in EEG recordings.

Detecting a peak indicative of a particular event in the EEG signal is challenging due to the non-stationary nature of the signal relative to the baseline amplitude, time, and different user. A signal peak identified as a point of highest signal amplitude lying between

two associated valley points, which hold a local minimum value. Any single peak is described by a number of signal parameters, including amplitude, width, and slope. Based on those parameters, a number of peak features can be calculated in the temporal domain, such as peak-to-peak amplitudes at the first half wave, peak width, ascending peak slopes at the first half wave, and descending peak slope at the second half wave. The ensemble of these peak features serves to detect the peak in various applications. However, the high variation of calculated peak features in real EEG data, which typically contain several types of noise, can interfere with correct peak detection and degrade performance.

Typically, a peak detection algorithm consists of combination of selected features from their peak model and subsequent computational processes, such as classification. Based on the literature, Dumpala et al. (1982) used a defined peak model, and then introduced a classification process to detect a peak signal in the analysis of gastric electrical activity. Dingle et al. (1993); Liu et al. (2002); Acir et al. (2005); Acir (2005); Liu et al. (2013) also used the defined peak models and different classification processes to detect peaks in EEG signal with epileptiform activity. The classifiers that have hitherto been used for signal peak detection include rule-based (Dumpala et al. 1982; Adam et al. 2015; Dingle et al. 1993), AdaBoost (Liu et al. 2002), radial basis function network (Acir et al. 2005), support vector machine (SVM) (Liu et al. 2013), radial basis support vector machine (Acir et al. 2005), artificial neural network (ANN) (Liu et al. 2002), and expert system (Liu et al. 2002; Dingle et al. 1993). In general, utilization of a peak detection algorithm provides the best performance in various applications. However, the various algorithms have used different peak models and different classification approaches combined in a particular peak detection algorithm. Moreover, to the best of our knowledge, there are very reports comparing the performance of different peak models using the same classification platform, such as a rule-based classifier (Adam et al. 2014a, 2014b, 2015). The existing methods tend to have poor performance with peak models defining many peak features, for example, the detection performance declined to nil when the classifier employed all 11 features from Liu model (Adam et al. 2014b).

For fair evaluation of the detection performance of different EEG signal peak models, they must be assessed using a common classification method. Therefore, in this study we used the extreme learning machine (ELM) method as a common classifier for the peak detection algorithm to evaluate the performance in association with four different peak models in time domain analysis, namely the models of Dumpala, Acir, Liu, and Dingle. The four representative peak models were on the basis of their proven utility in various physiological signal applications. We used an ELM since it provides very fast learning speed, generalized performance, learning without iterative tuning, and minimal requirement for user intervention. ELM also employs as alternative method to resolve the shortcoming of existing studies which have poor performance to peak model with many peak features. Hence, the present study aims to determine the best peak model in time domain analysis for EEG-based horizontal eye movement signal application using an advantageous common classification platform.

Methods

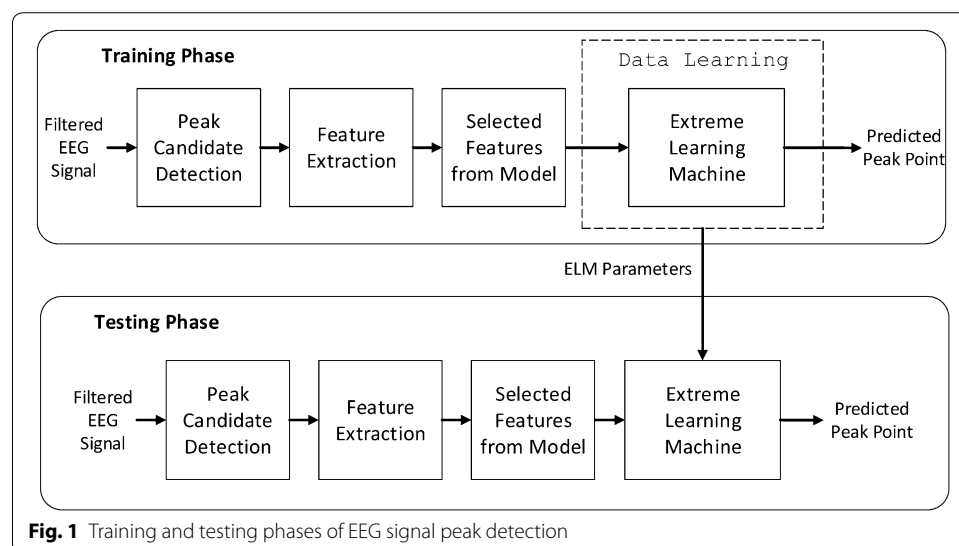
EEG signal peak detection algorithm

The training and testing phases of the EEG signal peak detection algorithm are shown in Fig. 1. The training and testing data that used in this study were collected using two channel EEG recordings from 20 voluntary subjects. In the first stage of peak detection, the training and testing EEG signals must be filtered as input to the algorithm, upon selection of the desired peak model. The training phase of the algorithm involves several processes, namely including peak candidate detection, feature extraction, with definition of model-specific features, and then classification process. The estimation process is performed during this phase to train the network for adjusting the ELM parameters using the learning algorithm of the ELM classifier. In the testing phase, the algorithm follows the same series of processes, and the ELM parameters first determined in the training phase are used in the classification process of the testing phase. The final output of the training and testing phase are the predicted peak points and non-peak points from the identified peak candidates.

Peak candidate detection

The first process of the detection algorithm is to determine candidate peaks. This process is used to assign the peaks into two groups, true non-peaks and true peaks. The group of true peaks is reconsidered to consist of candidate peaks, which can further classified into true non-peaks and true peaks using the ELM classifier. One advantage of determining candidate peaks process is that it reduces the number of input samples required for the ELM classifier, such that the computational time in the training and testing phases is minimized.

Determining the local maxima (peak points) and minima (valley points), the first process in determining a candidate peak, can be performed using an algorithm developed by Billauer (2012). The subsequent process of detecting a peak candidate is as follows: By considering a discrete-time signal, $x(I)$ of L points, the i th candidate peak point, PP_i , is identified using the three-point sliding window method (Dumpala et al. 1982). The three



selected points are denoted as $x(I-1)$, $x(I)$ and $x(I+1)$ for $I = 1, 2, 3, \dots, L$. A candidate peak point is identified when $x(PP_i-1) < x(PP_i) > x(PP_i+1)$ and two associated valley points, $VP1_i$ and $VP2_i$ lie on either side of the peak, as shown in Fig. 2. The valley points are defined when $x(VP1_i-1) > x(VP1_i) < x(VP1_i+1)$ and $x(VP2_i-1) > x(VP2_i) < x(VP2_i+1)$.

Feature extraction

The features of a peak candidate are calculated based on the eight points shown in Fig. 2. The set of points consists of the i th candidate peak point, PP_i , the two associated valley points, $VP1_i$ and $VP2_i$, the half point at first half wave ($HP1_i$), the half point at second half wave ($HP2_i$), the turning point at first half wave ($TP1_i$), the turning point at second half wave ($TP2_i$), and the moving average curve point ($MAC(PP_i)$). The half point at first half wave can be defined as the point in slope located in the middle between the PP_i and $VP1_i$ points while the half point at the second half wave as the point in slope based in the midst between the PP_i and $VP2_i$ points. The $MAC(PP_i)$ point is located at the intersection between the PP_i and $MAC(PP_i)$ points. The window length of the moving averaging is 100 sampling points.

Based on signal parameters, the features of a peak candidate can be categorized into three groups, namely amplitude, width, and slope. There are five different amplitudes, seven different widths, and four different slopes that can be calculated based on the eight defined points, resulting in a total of 16 features, which can be defined as follows:

1. The peak-to-peak amplitude at the first half wave, f_1 , is the peak amplitude between the magnitude of the peak and the magnitude of the valley of the first half wave, as denoted by.

$$f_1 = |y(PP_i) - y(VP1_i)| \quad (1)$$

2. The peak-to-peak amplitude at the second half wave, f_2 , is the peak amplitude between the magnitude of the peak and the magnitude of the valley of the second half wave, and is defined as

$$f_2 = |y(PP_i) - y(VP2_i)| \quad (2)$$

3. The turning point amplitude at the first half wave, f_3 , is the peak amplitude between the magnitude of the peak and the magnitude of the turning point at the first half-

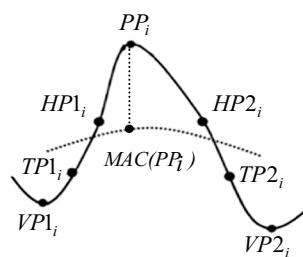


Fig. 2 The eight points of peak model

wave. The turning point is defined as the point where the slope decreases more than 50 % compared to the slope of the preceding point. The equation for f_3 is as follows:

$$f_3 = |y(PP_i) - y(TP1_i)| \quad (3)$$

4. The turning point amplitude at the second half wave, f_4 , is the peak amplitude between the magnitude of the peak and the magnitude of the turning point at the second half wave, and is defined as

$$f_4 = |y(PP_i) - y(TP2_i)| \quad (4)$$

5. The moving average amplitude, f_5 , is the peak amplitude between the magnitude of the peak and the magnitude of the moving average, and is defined as

$$f_5 = |y(PP_i) - y(MAC(PP_i))| \quad (5)$$

6. The peak width, f_6 , is the peak width between the valley point of the first half point and the valley point of the second half wave, and is defined as

$$f_6 = |x(VP1_i) - x(VP2_i)| \quad (6)$$

7. The first half wave width, f_7 , is the peak width between the peak point and the valley point of the first half wave, and is defined as

$$f_7 = |x(PP_i) - x(VP1_i)| \quad (7)$$

8. The second half wave width, f_8 , is the peak width between the peak point and the valley point of the second half wave, and is defined as

$$f_8 = |x(PP_i) - x(VP2_i)| \quad (8)$$

9. The turning point width, f_9 , is the peak width between the turning point at the first half wave and the turning point at the second half wave, and is defined as

$$f_9 = |x(TP1_i) - x(TP2_i)| \quad (9)$$

10. The first half-wave turning point width, f_{10} , is the peak width between the turning point at the first half-wave and the peak point, and is defined as

$$f_{10} = |x(PP_i) - x(TP1_i)| \quad (10)$$

11. The second half wave turning point width, f_{11} , is the peak width between the turning point at the second half-wave and the peak point, and is defined as

$$f_{11} = |x(PP_i) - x(TP2_i)| \quad (11)$$

12. The half point width, f_{12} , is the peak width between the half point of the first half-wave and the half point of the second half-wave, and is defined as

$$f_{12} = |x(HP1_i) - x(HP2_i)| \quad (12)$$

13. The peak slope at the first half wave, f_{13} , is the maximal slope between the peak point and the valley point of the first half wave, and is defined as

$$f_{13} = \left| \frac{y(PP_i) - y(VP1_i)}{x(PP_i) - x(VP1_i)} \right| \quad (13)$$

14. The peak slope at the second half-wave, f_{14} , is the peak slope between the peak point and the valley point of the second half wave, and is defined as

$$f_{14} = \left| \frac{y(PP_i) - y(VP2_i)}{x(PP_i) - x(VP2_i)} \right| \quad (14)$$

15. The turning point slope at the first half-wave, f_{15} , is the peak slope between the peak point and the turning point of the first half-wave, and is defined as

$$f_{15} = \left| \frac{y(PP_i) - y(TP1_i)}{x(PP_i) - x(TP1_i)} \right| \quad (15)$$

16. The turning point slope at the second half wave, f_{16} , is the peak slope between the peak point and the turning point of the second half-wave, and is defined as

$$f_{16} = \left| \frac{y(PP_i) - y(TP2_i)}{x(PP_i) - x(TP2_i)} \right| \quad (16)$$

From these 16 features, Dumpala et al. (1982) introduced a peak model that uses the four most salient features, f_1, f_6, f_{13} , and f_{14} . Additional features defining peak amplitude, f_2 , and two features of peak width, f_7 and f_8 , were introduced by Acir et al. (2005); Acir and Guzeli (2004). As pointed out in Acir et al. (2005), the defined features are inter-related to the characteristic of peak in epilepsy events. There are highlighted three characteristics as follows; (1) the ascending peak slopes at the first half wave, and descending peak slope at the second half wave are relatively large and smooth (2) the top of the peak is sharp, and (3) the peak width is always between 20 and 70 ms. The peak width can be calculated by the sum of the first half wave and second half wave peak width. Dumpala et al. (1982) and Acir et al. (2005) used the same similar definition of the peak slopes, i.e. f_{13} and f_{14} . The peak model of Liu et al. (2002) entailed a total of 11 features, consisting of four amplitudes (f_1, f_2, f_3 , and f_4), three widths (f_6, f_9, f_{12}), and four slopes ($f_{13}, f_{14}, f_{15}, f_{16}$). Finally, the peak model introduced by Dingle et al. (1993) consists of four features (f_5, f_6, f_{13}, f_{14}). The different peak models and their sets of features are listed in Table 1.

Extreme learning machine classifier

Extreme learning machine is a new approach to machine learning involving a single layer feedforward neural network (SLFN). The introduction of ELM by Huang et al. (2004) has been followed by several variants (Balasundaram et al. 2014; Huang et al. 2012). ELM has already been used with success for various event classifications of EEG signals (Song and Zhang 2013; Shi and Lu 2013; Yuan et al. 2012; Song et al. 2012; Yuan et al. 2011; Liang et al. 2006). The main advantage present by ELM is the fast computation of the learning

Table 1 List of different peak models and sets of features

Peak model	Set of features	Number of features
Dumpala et al. (1982)	f_1, f_6, f_{13}, f_{14}	4
Dingle et al. (1993)	f_5, f_6, f_{13}, f_{14}	4
Acir et al. (2005); Acir (2005); Acir and Guzeli (2004)	$f_1, f_2, f_7, f_8, f_{13}, f_{14}$	6
Liu et al. (2002)	$f_1, f_2, f_3, f_4, f_6, f_9, f_{12}, f_{13}, f_{14}, f_{15}, f_{16}$	11

method compared to conventional ANN learning, since ELM training dispenses with time-consuming iterative tuning.

The architecture of an ELM is shown in Fig. 3. The network consists of three layers, i.e. the input, hidden, and output layers. Between the input and hidden layers are the input weights, and between the hidden and output layers are the output weights. The training process of an ELM proceeds in three stages. In the first stage, the input weights are assigned randomly between -1 and 1 , and the biases in the hidden layer are assigned randomly between 0 and 1 . Both of these parameters remain fixed during the training process. Afterward, the output matrix of the hidden layer, H , is calculated as follows:

$$H = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_N) \end{bmatrix} = \begin{bmatrix} g\left(\sum_{i=1}^d a_{i1}x_{1i} + b_1\right) & \cdots & g\left(\sum_{i=1}^d a_{iL}x_{1i} + b_L\right) \\ \vdots & \ddots & \vdots \\ g\left(\sum_{i=1}^d a_{i1}x_{Ni} + b_1\right) & \cdots & g\left(\sum_{i=1}^d a_{iL}x_{Ni} + b_L\right) \end{bmatrix}_{N \times L} \quad (17)$$

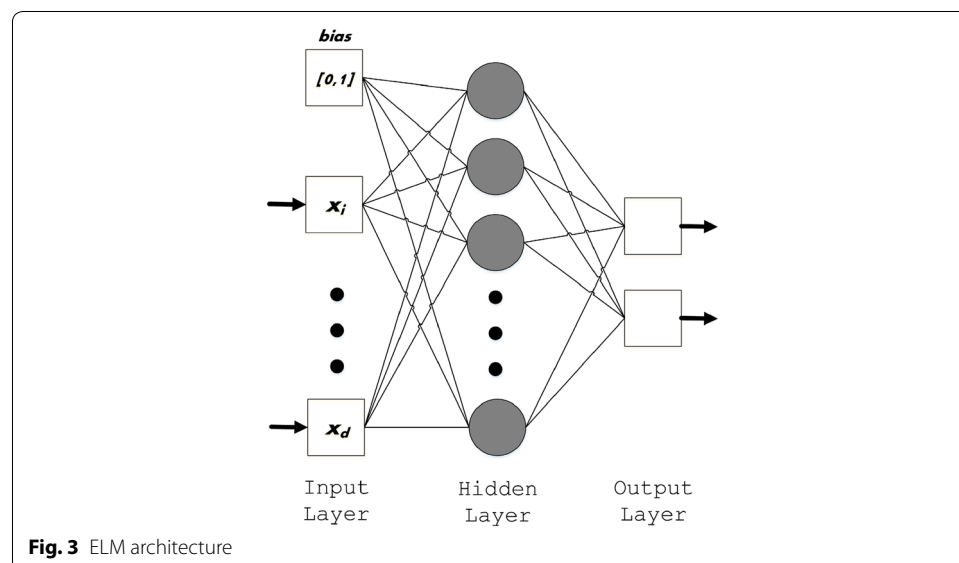
where g is an activation function of the hidden neuron, x is the $N \times L$ matrix of inputs, a is the $d \times L$ matrix of random input weights, b is the $1 \times L$ matrix of random biases in the hidden layer, N is an arbitrary number of distinct samples, L is the number of hidden neurons, and d is the number of inputs or features. The i th column of H is the output of the i th hidden neuron with respect to inputs x_1, x_2 , until x_d .

The ELM can be represented as a linear system, which is mathematically modeled as

$$H\beta = T \quad (18)$$

where β is the $L \times m$ matrix of output weights, T is the $N \times m$ matrix of target outputs, and m is the number of output neurons. The β and T matrixes are denoted as

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times m} \quad (19)$$



and

$$T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times m} \quad (20)$$

respectively. To determine the least square solution, β , of the linear system $H\beta = T$, the minimum norm least-squares solution is computed as follows:

$$\|H(a_1, \dots, a_L, b_1, \dots, b_L)\beta - T\| = \min_{\beta} \|H(a_1, \dots, a_L, b_1, \dots, b_L)\beta - T\| \quad (21)$$

It is well known that the smallest norm least-squares solution of Eq. 21 is

$$\beta = (H^T H)^{-1} H^T T = H^+ T \quad (22)$$

where H^+ is the Moore–Penrose pseudo-inverse of H . The three training stages of the ELM classifier are summarized as follows:

- Stage 1: Randomly assign the input weights, a_i and biases in the hidden neurons, b_i
- Stage 2: Calculate the output matrix of the hidden layer, H
- Stage 3: Calculate the output weights, $\beta = H^+ T$.

The output function of the ELM classifier of a given unknown sample x is

$$f(x) = h(x)\beta \quad (23)$$

In the output layer, two neurons are used in the network to classify the output into two classes (output): class 1 and class 0. For the two classes ($m > 1$), the predicted class label is the i th number of the output neurons, which is the maximum value of the output neuron (Huang et al. 2012). The predicted class label of a given unknown sample x is defined as follows.

$$label(x) = \arg \max_{i \in \{1, \dots, m\}} f_i(x) \quad (24)$$

We evaluate the performance of our ELM classifier based on the G_{mean} (Guo et al. 2008), which is calculated as follows:

$$TPR = \frac{TP}{TP + FN} \quad (25)$$

$$TNR = \frac{TN}{TN + FP} \quad (26)$$

$$G_{mean} = \sqrt{TPR \times TNR} \quad (27)$$

where a true peak (TP) is the correctly-detected peak point of the peak candidate, a true non-peak (TN) is the correctly-detected non-peak point of the peak candidate, a false peak (FP) is the wrongly-detected non-peak point of a peak candidate, a false non-peak

(*FN*) is a wrongly-detected peak point of a peak candidate, *TPR* is the true peak rate, and *TNR* is the true non-peak rate.

Experimental setup and protocols

Each experiment is conducted in 30 independent runs. The first 50 % of the filtered EEG signal was used as training data, and the remaining 50 % as testing data.

The parameter settings of the ELM classifier are shown in Table 2. The number of hidden neuron was selected using a trial and error method, which was set to 500. The sigmoid [−1, 1] was used as an activation function in the hidden layer for the purpose of normalization, whereas a linear function was located on each neuron in the output layer. Other settings for the ELM classifier, such as the number of neurons in the input layer are dependent on the number of selected features of a particular peak model. The number of output neurons was set to 2 which it used maximum argument as indicated in Eq. 24 for choosing the ELM output. We note that the input weights and the biases remained fixed during the training, but the values of these two ELM parameters are randomly assigned for each of 30 runs.

This experimental protocol was approved by the medical ethics committee of the University of Malaya Medical Centre. All subjects signed informed consent forms in advance.

The filtered EEG signals in this study were obtained in the Applied Control and Robotic (ACR) Laboratory, Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Malaysia. Twenty healthy subjects (10 males and 10 females, aged 20–40 years), who are undergraduate and postgraduate students in the Faculty of Engineering, volunteered to participate in these data collection sessions. Filtered EEG signal recordings were obtained using the g.MOBILab portable biological signal acquisition system. The scalp electrodes were arranged using the 10–20 international electrode placement system. The EEG signal was recorded from the C3 and C4 channels, with the signal of channel CZ used as a reference. The ground electrode was located on the FPz channel, such that a total of only four electrodes were used. The sampling frequency was set to 256 Hz. The electrodes from the C3 and C4 channels are positioned for detecting EEG peaks associated with the brain response of commanded horizontal eye gaze direction. We used the C3, C4, and CZ channels are used because of they have relatively little less contamination from EEG artifacts due to eye blinking (Klados et al. 2011).

All subjects had been instructed to get a good rest the night before the data collection session, so as to ensure full focus during the EEG recordings. The subjects were told to

Table 2 Parameter settings of ELM

Parameters	Value
Number of neurons in the hidden layer	500
Biases in the hidden layer	Random [0, 1]
Activation function in the hidden layer	Sigmoid [−1, 1]
Activation function in the output layer	Linear function
Number of neurons in the input layer	Depends on a number of features
Number of neurons in the output layer	2

prepare for the external voice cue within up to 4 s. Appearance of the cue is voice command or verbal reminder for the subject to move his eyes initially forward fixation to the left or to the right. At exactly 5 s from the beginning session, the external voice cue appears randomly instructing the subject to shift gaze to the left or right direction, and hold the new eye position from 5 until 10 s, which is the end of the EEG recording. The eye gaze directions that produce a number of peaks in the signal on channels C3 and C4 are archived as raw data for analysis.

Figure 4 shows a representative case of filtered EEG signals that are labeled as eye movement signals. The dotted red vertical lines show the actual peak point locations, as assigned by a researcher. The eye movement signal consists of 20 signals for channel C3, 20 signals for channel C4, for duration of 10-s per signal, recorded at 256 Hz for a total of 2560 sampling points per signal. Furthermore, each signal contains one known peak point location, where the known peak pattern represents the eye gaze direction, either to the left or to the right.

Experimental results and discussion

Four different outcome measures were used in the experiments: the average G_{mean} , the maximum G_{mean} , the minimum G_{mean} , and the standard deviation (STDEV). Additionally, the statistical comparisons of the average test accuracy among the four models are evaluated using Friedman's test.

The results of training and testing performance based on the four different measurements (end-points) for the peak model with all 16 features are shown in Table 3. For training performance, the average, maximum, minimum, and STDEV values were 88.3, 94.9, 80.6, and 3.6 %, respectively, whereas for testing performance, the average, maximum, minimum, and STDEV values were 36.9, 58.1, 0, and 11.8 %, respectively. The minimum testing performance of 0 % indicates that the classifier was unable to correctly detect even one peak.

Next, the training and testing performance based on the four different measurements for each peak model are shown in Table 4, the training performance average, maximum, minimum, and STDEV values are 84.7, 86.6, 83.7 and 1.4 % for Dumpala peak model,

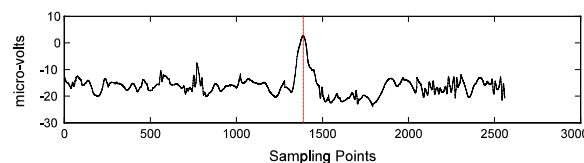


Fig. 4 Filtered EEG-based eye movement signal (one peak point per signal)

Table 3 Peak detection training and testing performance using all features

Measurement	Training (%)	Testing (%)
Average	88.3	36.9
Maximum	94.9	58.1
Minimum	80.6	0
STDEV	3.6	11.8

Table 4 Peak detection training and testing performance for each peak model

Peak model	Training (%)				Testing (%)			
	Average	Max	Min	SD	Average	Max	Min	SD
Dumpala	84.7	86.6	83.7	1.4	70.1	82.6	51.6	6.7
Acir	88.3	89.4	86.6	1.4	36.9	62.6	0	11.9
Liu	78.9	83.7	74.1	2.6	52.1	71.8	37.2	7.9
Dingle	99.5	100	97.4	0.9	71.7	89.2	57.1	6.9

whereas the corresponding results were 88.3, 89.4, 86.6 and 1.4 % for Acir peak model, 78.9, 83.7, 74.1 and 2.6 % for Liu peak model, and 99.5, 100, 97.4 and 0.9 %, for Dingle peak model. The performance of Dingle peak model was clearly superior to that of the other peak models.

The testing performance on average, maximum, minimum, and STDEV values are 70.1, 82.6, 51.6 and 6.7 %, respectively, for Dumpala peak model; 36.9, 62.6, 0 and 11.9 %, respectively, for Acir peak model; 51.1, 71.8, 37.2 and 7.9 %, respectively, for Liu peak model and 71.7, 89.2, 57.1 and 6.9 %, respectively, for Dingle peak model. As with the training data, performance of Dingle peak model with the test data was superior to that of the other peak models. As with the training set, results with Acir method for the test data showed that the classifier is unable to correctly predict all the true peaks occurring at the particular time when TP is equal to zero. Thus, G_{mean} becomes zero even though the TN value is non-zero.

Sensitivity is the percentage of true peak rate recovered while specificity is the percentage of true non-peak rate. The overall sensitivity and specificity values for the testing performance are shown in Table 5. The results in Table 5 show that sensitivity is significantly lower than 30 % for the Acir and Liu models. Dumpala and Dingle peak models performed best, with sensitivity of 55 % and specificity exceeding 99 %. Overall, the sensitivity of the four peak models is lower than their specificity, thus resulting in a large amount of false non-peak. The four peak models return many false non-peaks due to several contributing factors such the collected data is affected by various noises and the peak features have a large different value from one subject to another subject. These factors are the cause to the high variation of peak features of the four peak models. In this case, the NNRW classifier has performed best for classifying the non-peak features than peak features.

The comparison of the average test accuracy between the four peak models is extended using Friedman's test for statistical analysis. The analysis searches for a significant difference in the average testing accuracies between the peak models with p value lower than

Table 5 Sensitivity and specificity testing performance for each peak model

Peak model	Sensitivity (%) = $\frac{TP}{TP+FN} \times 100$	Specificity (%) = $\frac{TN}{TN+FP} \times 100$
Dumpala	57.2	99.6
Acir	18	99.9
Liu	28.3	99.7
Dingle	55	99.7

threshold of 0.01. The average rankings of Friedman's test (Table 6) show best results for Dingle peak model, followed by the Dumpala, Liu, and Acir peak models. Post-hoc analysis of Friedman's test results are based on the Holm-Bonferroni method using two difference confidence intervals, $\alpha = 0.05$ and $\alpha = 0.10$ as shown in Table 7. Both Friedman's test and the Holm-Bonferroni post hoc analysis are conducted using the KEEL software tool (Alcala-Fdez et al. 2009). The post hoc results in Table 7 show similar rank orders for $\alpha = 0.05$ and $\alpha = 0.10$, where Dingle peak model offers test accuracies significantly better than do Acir and Liu peak models, but was not superior to without any significant compared to Dumpala peak model.

Conclusions and future work

In this study, we applied ELM-based peak detection to two-lead EEG signals recorded from 20 healthy subjects instructed to direct their horizontal gaze in response to a voice cue. The data was used to evaluate the performance of four different peak detection models. The various event-related EEG peaks were analyzed through a series of processes, i.e. peak candidate detection, feature extraction, and classification. The four peak models considered in this study are representative of typical EEG studies in the literature (Dumpala et al. 1982; Acir et al. 2005; Acir 2005; Dingle et al. 1993; Liu et al. 2002), all of which entail initial extraction of 16 peak features. Each of the four peak models was selected in turn before the execution of experiments using the ELM as a common classification algorithm. ELM has been tested on more than forty benchmark data sets. Also, ELM has been proven by experimental results that achieved similar or better generalized performance compared to SVM and least square support vector machine (LS-SVM) for two-class classification (Huang et al. 2012). We find Dingle peak model to be the best for reliably detecting voluntary horizontal eye movement signal peaks, delivering a mean performance of 99.5 % for the training set and 71.7 % for the testing set. The testing performance needs to be improved by reconsidering the selection of peak features among

Table 6 Average ranking of Friedman's test with $p < 0.01$

Peak model	Average ranking	Rank
Dumpala	1.5667	2
Acir	3.7	4
Liu	3.2333	3
Dingle	1.5	1

Table 7 Post-hoc analysis for Friedman's test

<i>i</i>	Condition	$\alpha = 0.05$		$\alpha = 0.10$	
		<i>p</i>	Holm	<i>p</i>	Holm
6	Acir vs. Dingle	0.000001	0.00833	0.000001	0.01667
5	Dumpala vs. Acir	0.000001	0.01	0.000001	0.02
4	Liu vs. Dingle	0.000001	0.0125	0.000001	0.025
3	Dumpala vs. Liu	0.000001	0.01667	0.000001	0.03333
2	Acir vs. Liu	0.161513	0.025	0.161513	0.05
1	Dumpala vs. Dingle	0.841481	0.05	0.841481	0.1

16 features and exploring other variant of ELM classifiers. Furthermore, the results in Table 4 also indicate that Dingle peak model to be a good generalized model, as revealed by the highest classification rate of the minimum testing result at 57.1 %. Additionally, Friedman's test confirms that Dingle peak model offers significant better average test accuracy than those of Acir and Liu models.

This study also observes that defining more peak features on model is not guarantee in producing better accuracy on EEG-based horizontal eye movement signal application. As shown in the results in Table 3, the mean of testing accuracy only can achieve at 36.9 %. However, determining the optimal model from the selected features associated with the advantageous of common classification platform is the best approach to gain the accuracy of detection performance.

Results of this study may be applicable in many contexts characterized by the general problem of signal detection, including, such as medical diagnostics, human-machine interface (HMI), brain-computer interface (BCI), and harmonic detection in digital and audio signal processing. For example, an EEG peak in the frontal eye field associated with a change of horizontal eye gaze direction could be translated to the direction of cursor movement in BCI applications, which might be useful for patients with locked-in syndrome or other disabilities (Belkacem et al. 2014). This approach might also be translatable for EEG-based command of the movement of a robotic arm or wheelchair in HMI applications (Postelnicu et al. 2011; Ramli et al. 2015). We intend in the future to extend this work to the problem of feature selection for the peak detection algorithm, so as to optimize the selection of the most salient peak features, with an aim to improve further the performance of peak detection.

Authors' contributions

AA conceived the study, participated in the design of the algorithm, carries out collected the data, conducted experiments, performed the statistical analysis, and drafted the manuscript. ZI participated in the design of the study, coordination and helped to draft the manuscript. NM prepared the facilities in the laboratory, financing, and participated in the design of the study. MIS contributed to the financing publication fee, the design of the study, manuscript preparation, manuscript editing, and the experiments facilities. PC contributed to the manuscript editing. MB contributed to the facilities in the laboratory. All authors read and approved the final manuscript.

Author details

¹ Applied Control and Robotics (ACR) Laboratory, Department of Electrical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia. ² Faculty of Electrical and Electronic Engineering, Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia. ³ Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia Kuala Lumpur, Jalan Semarak, 54100 Kuala Lumpur, Malaysia. ⁴ School of Psychology and Counseling, Queensland University of Technology, and QIMR Berghofer, Brisbane, Australia.

Acknowledgements

This research is funded by High Impact Research Fund (UM.C/HIR/MOHE/ENG/16 Account code: D000016-16001), awarded by Ministry of Education Malaysia to University of Malaya. Also thanks to Universiti Teknologi Malaysia for funding this research work through Research University Grant (GUP) (Q.K130000.2643.10J98). The first author would like to thank the Ministry of Education Malaysia for supporting his study by awarding him a MyPhD scholarship.

Competing interests

The authors declare that they have no competing interests.

Ethics approval and consent to participate

All experimental protocols performed in this study involving data from human participations were approved by the medical ethics committee of the University of Malaya Medical Centre. All subjects signed informed consent forms in advance.

Received: 10 March 2016 Accepted: 27 June 2016

Published online: 11 July 2016

References

- Acir N (2005) Automated system for detection of epileptiform patterns in EEG by using a modified RBFN classifier. *Expert Syst Appl* 29(2):455–462. doi:[10.1016/j.eswa.2005.04.040](https://doi.org/10.1016/j.eswa.2005.04.040)
- Acir N, Guzel C (2004) Automatic spike detection in EEG by a two-stage procedure based on support vector machines. *Comput Biol Med* 34(7):561–575. doi:[10.1016/j.combiomed.2003.08.003](https://doi.org/10.1016/j.combiomed.2003.08.003)
- Acir N, Kuntalp IO, Baklan B, Güzelis C (2005) Automatic detection of epileptiform events in EEG by a three-stage procedure based on artificial neural networks. *IEEE Trans Biomed Eng* 52(1):30–40. doi:[10.1109/TBME.2004.839630](https://doi.org/10.1109/TBME.2004.839630)
- Adam A, Ibrahim Z, Mokhtar N, Shapiai MI, Mohd Tumari MZ, Mubin M (2014a) Feature selection and classifier parameter estimation for EEG signal peak detection using gravitational search algorithm. In: 4th international conference on artificial intelligence with applications in engineering and technology, IEEE, Kota Kinabalu, 3–5 Dec 2014a. pp 103–108
- Adam A, Shapiai MI, Mohd Tumari MZ, Mohamad MS, Mubin M (2014b) Feature selection and classifier parameters estimation for EEG signals peak detection using particle swarm optimization. *Sci World J* (Article ID 973063):1–13. doi:[10.1155/2014/973063](https://doi.org/10.1155/2014/973063)
- Adam A, Ibrahim Z, Mokhtar N, Shapiai MI, Mubin M (2015) Dingle's model-based EEG peak detection using a rule-based classifier. In: 2015 International conference on artificial life and robotics, Oita, 10–12 Jan 2015. p 53
- Alcala-Fdez J, Sanchez L, Garcia S, del Jesus MJ, Ventura S, Garrell JM, Otero J, Romero C, Bacardit J, Rivas VM, Fernandez JC, Herrera F (2009) KEEL: a software tool to assess evolutionary algorithms for data mining problems. *Soft Comput* 13(3):307–318. doi:[10.1007/s00500-008-0323-y](https://doi.org/10.1007/s00500-008-0323-y)
- Balasundaram S, Gupta D, Kapil (2014) 1-Norm extreme learning machine for regression and multiclass classification using Newton method. *Neurocomputing* 128:4–14. doi:[10.1016/j.neucom.2013.03.051](https://doi.org/10.1016/j.neucom.2013.03.051)
- Barea R, Boquete L, Ortega S, Lopez E, Rodriguez-Ascariz JM (2012) EOG-based eye movements codification for human computer interaction. *Expert Syst Appl* 39(3):2677–2683. doi:[10.1016/j.eswa.2011.08.123](https://doi.org/10.1016/j.eswa.2011.08.123)
- Belkacem AN, Hirose H, Yoshimura N, Shin D, Koike Y (2014) Classification of four eye directions from EEG signals for eye-movement-based communication systems. *J Med Biol Eng* 34(6):581–588. doi:[10.5405/Jmbe.1596](https://doi.org/10.5405/Jmbe.1596)
- Dingle AA, Jones RD, Carroll G, Fright WR (1993) A multistage system to detect epileptiform activity in the EEG. *IEEE Trans Biomed Eng* 40(12):1260–1268. doi:[10.1109/10.250582](https://doi.org/10.1109/10.250582)
- Dumpala SR, Reddy SN, Sarna SK (1982) An algorithm for the detection of peaks in biological signals. *Comput Prog Biomed* 14(3):249–256. doi:[10.1016/0010-468X\(82\)90030-7](https://doi.org/10.1016/0010-468X(82)90030-7)
- Elgendi M, Norton I, Brearley M, Abbott D, Schuurmans D (2013) Systolic peak detection in acceleration photoplethysmograms measured from emergency responders in tropical conditions. *PLoS One* 8(10):e76585. doi:[10.1371/journal.pone.0076585](https://doi.org/10.1371/journal.pone.0076585)
- Guo X, Yin Y, Dong C, Yang G, Zhou G (2008) On the Class imbalance problem. Paper presented at the proceedings of the fourth international conference on natural computation (ICNC 08), Jinan, 25–27 Aug 2008
- Huang GB, Zhu QY, Siew CK (2004) Extreme learning machine: a new learning scheme of feedforward neural networks. In: 2004 IEEE international joint conference on neural networks, 25–29 July 2004. pp 985–990
- Huang GB, Zhou HM, Ding XJ, Zhang R (2012) Extreme learning machine for regression and multiclass classification. *IEEE T Syst Man Cy B* 42(2):513–529. doi:[10.1109/Tsmcb.2011.2168604](https://doi.org/10.1109/Tsmcb.2011.2168604)
- Klados MA, Papadelis C, Braun C, Bamidis PD (2011) REG-ICA: a hybrid methodology combining blind source separation and regression techniques for the rejection of ocular artifacts. *Biomed Signal Process Control* 6(3):291–300. doi:[10.1016/j.bspc.2011.02.001](https://doi.org/10.1016/j.bspc.2011.02.001)
- Liang NY, Saratchandran P, Huang GB, Sundararajan N (2006) Classification of mental tasks from EEG signals using extreme learning machine. *Int J Neural Syst* 16(1):29–38. doi:[10.1142/s0129065706000482](https://doi.org/10.1142/s0129065706000482)
- Liu HS, Zhang T, Yang FS (2002) A multistage, multimethod approach for automatic detection and classification of epileptiform EEG. *IEEE Trans Biomed Eng* 49(12):1557–1566. doi:[10.1109/TBME.2002.805477](https://doi.org/10.1109/TBME.2002.805477)
- Liu Y-C, Lin C-CK, Tsai J-J, Sun Y-N (2013) Model-based spike detection of epileptic EEG data. *Sensors* 13(9):12536–12547. doi:[10.3390/s130912536](https://doi.org/10.3390/s130912536)
- Peakdet Peak detection using MATLAB (2012) <http://www.billauer.co.il/peakdet.html>
- Postelnicu CC, Talaba D, Toma MI (2011) Controlling a robotic arm by brainwaves and eye movement. In: technological innovation for sustainability, vol 349. IFIP Advances in Information and Communication Technology. pp 157–164. doi:[10.1007/978-3-642-19170-1_17](https://doi.org/10.1007/978-3-642-19170-1_17)
- Ramli R, Arof H, Ibrahim F, Mokhtar N, Idris MYI (2015) Using finite state machine and a hybrid of EEG signal and EOG artifacts for an asynchronous wheelchair navigation. *Expert Syst Appl* 42(5):2451–2463. doi:[10.1016/j.eswa.2014.10.052](https://doi.org/10.1016/j.eswa.2014.10.052)
- Shi L-C, Lu B-L (2013) EEG-based vigilance estimation using extreme learning machines. *Neurocomputing* 102:135–143. doi:[10.1016/j.neucom.2012.02.041](https://doi.org/10.1016/j.neucom.2012.02.041)
- Song Y, Zhang J (2013) Automatic recognition of epileptic EEG patterns via extreme learning machine and multiresolution feature extraction. *Expert Syst Appl* 40(14):5477–5489. doi:[10.1016/j.eswa.2013.04.025](https://doi.org/10.1016/j.eswa.2013.04.025)
- Song Y, Crowcroft J, Zhang J (2012) Automatic epileptic seizure detection in EEGs based on optimized sample entropy and extreme learning machine. *J Neurosci Methods* 210(2):132–146. doi:[10.1016/j.jneumeth.2012.07.003](https://doi.org/10.1016/j.jneumeth.2012.07.003)
- Yuan Q, Zhou W, Li S, Cai D (2011) Epileptic EEG classification based on extreme learning machine and nonlinear features. *Epilepsy Res* 96(1–2):29–38. doi:[10.1016/j.epilepsyres.2011.04.013](https://doi.org/10.1016/j.epilepsyres.2011.04.013)
- Yuan Q, Zhou W, Zhang J, Li S, Cai D, Zeng Y (2012) EEG classification approach based on the extreme learning machine and wavelet transform. *Clin Eeg Neurosci* 43(2):127–132. doi:[10.1177/1550059411435861](https://doi.org/10.1177/1550059411435861)